RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _____09 /057, 016_____

Source: _____1600_____

Date Processed by STIC: _____02-23-2005_____

ENTERED

RECEIVED

NOV 1 4 2005

TECH CENTER 1600/2900

1600

RAW SEQUENCE LISTING   DATE: 02/23/2005
PATENT APPLICATION: **US/09/057,016**   TIME: 15:06:06

Input Set : **N:\Crf3\RULE60\09057016.raw.txt**
Output Set: **N:\CRF4\02232005\I057016.raw**

SEQUENCE LISTING

```
 3 (1) GENERAL INFORMATION:
 5      (i) APPLICANT: Turpen, Thomas H.
 6                    Reinl, Stephen
 7                    Grill, Laurence K.
 9     (ii) TITLE OF INVENTION: Production of Peptides in Plants as
10                             Viral Coat Protein Fusions
12    (iii) NUMBER OF SEQUENCES: 27
14     (iv) CORRESPONDENCE ADDRESS:
15          (A) ADDRESSEE: Pennie & Edmonds
16          (B) STREET: 1155 Avenue of the Americas
17          (C) CITY: New York
18          (D) STATE: New York
19          (E) COUNTRY: USA
20          (F) ZIP: 10036
22      (v) COMPUTER READABLE FORM:
23          (A) MEDIUM TYPE: Floppy disk
24          (B) COMPUTER: IBM PC compatible
25          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
28     (vi) CURRENT APPLICATION DATA:
```
C--> 34 `         (A) APPLICATION NUMBER: US/09/057,016`
C--> 35 `         (B) FILING DATE: 07-Apr-1998`
```
31          (C) CLASSIFICATION: 435
33   (viii) ATTORNEY/AGENT INFORMATION:
36          (A) NAME: Halluin, Albert P.
37          (B) REGISTRATION NUMBER: 25,227
38          (C) REFERENCE/DOCKET NUMBER: 8129-087
40     (ix) TELECOMMUNICATION INFORMATION:
41          (A) TELEPHONE: 415-854-3660
42          (B) TELEFAX: 415-854-3694
43          (C) TELEX: 66141 PENNIE
46 (2) INFORMATION FOR SEQ ID NO: 1:
48      (i) SEQUENCE CHARACTERISTICS:
49          (A) LENGTH: 49 base pairs
50          (B) TYPE: nucleic acid
51          (C) STRANDEDNESS: unknown
52          (D) TOPOLOGY: unknown
54     (ii) MOLECULE TYPE: DNA (genomic)
58     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
60 GGAATTCAAG CTTAATACGA CTCACTATAG TATTTTTACA ACAATTACC              49
62 (2) INFORMATION FOR SEQ ID NO: 2:
64      (i) SEQUENCE CHARACTERISTICS:
```

RAW SEQUENCE LISTING   DATE: 02/23/2005
PATENT APPLICATION: US/09/057,016   TIME: 15:06:06

Input Set : N:\Crf3\RULE60\09057016.raw.txt
Output Set: N:\CRF4\02232005\I057016.raw

```
65            (A) LENGTH: 18 base pairs
66            (B) TYPE: nucleic acid
67            (C) STRANDEDNESS: unknown
68            (D) TOPOLOGY: unknown
70      (ii) MOLECULE TYPE: DNA (genomic)
74      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
76 CCTTCATGTA AACCTCTC                                          18
78 (2) INFORMATION FOR SEQ ID NO: 3:
80       (i) SEQUENCE CHARACTERISTICS:
81            (A) LENGTH: 25 base pairs
82            (B) TYPE: nucleic acid
83            (C) STRANDEDNESS: unknown
84            (D) TOPOLOGY: unknown
86      (ii) MOLECULE TYPE: DNA (genomic)
90      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
92 TAATCGATGA TGATTCGGAG GCTAC                                  25
94 (2) INFORMATION FOR SEQ ID NO: 4:
96       (i) SEQUENCE CHARACTERISTICS:
97            (A) LENGTH: 36 base pairs
98            (B) TYPE: nucleic acid
99            (C) STRANDEDNESS: unknown
100           (D) TOPOLOGY: unknown
102     (ii) MOLECULE TYPE: DNA (genomic)
106     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
108 AAAGTCTCTG TCTCCTGCAG GGAACCTAAC AGTTAC                     36
110 (2) INFORMATION FOR SEQ ID NO: 5:
112      (i) SEQUENCE CHARACTERISTICS:
113           (A) LENGTH: 36 base pairs
114           (B) TYPE: nucleic acid
115           (C) STRANDEDNESS: unknown
116           (D) TOPOLOGY: unknown
118     (ii) MOLECULE TYPE: DNA (genomic)
122     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
124 ATTATGCATC TTGACTACCT AGGTTGCAGG ACCAGA                     36
126 (2) INFORMATION FOR SEQ ID NO: 6:
128      (i) SEQUENCE CHARACTERISTICS:
129           (A) LENGTH: 24 base pairs
130           (B) TYPE: nucleic acid
131           (C) STRANDEDNESS: unknown
132           (D) TOPOLOGY: unknown
134     (ii) MOLECULE TYPE: DNA (genomic)
138     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
140 GGCGATCGGG CTGGTGACCG TGCA                                  24
142 (2) INFORMATION FOR SEQ ID NO: 7:
144      (i) SEQUENCE CHARACTERISTICS:
145           (A) LENGTH: 24 base pairs
146           (B) TYPE: nucleic acid
147           (C) STRANDEDNESS: unknown
148           (D) TOPOLOGY: unknown
```

RAW SEQUENCE LISTING   DATE: 02/23/2005
PATENT APPLICATION: US/09/057,016   TIME: 15:06:06

Input Set : N:\Crf3\RULE60\09057016.raw.txt
Output Set: N:\CRF4\02232005\I057016.raw

```
150     (ii) MOLECULE TYPE: DNA (genomic)
154     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
156 CGGTCACCAG CCCGATCGCC TGCA                                    24
158 (2) INFORMATION FOR SEQ ID NO: 8:
160     (i) SEQUENCE CHARACTERISTICS:
161         (A) LENGTH: 45 base pairs
162         (B) TYPE: nucleic acid
163         (C) STRANDEDNESS: unknown
164         (D) TOPOLOGY: unknown
166     (ii) MOLECULE TYPE: DNA (genomic)
170     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
172 CTAGCAATTA CAAGGTCCAG GTGCACCTCA AGGTCCTGGA GCTCC            45
174 (2) INFORMATION FOR SEQ ID NO: 9:
176     (i) SEQUENCE CHARACTERISTICS:
177         (A) LENGTH: 45 base pairs
178         (B) TYPE: nucleic acid
179         (C) STRANDEDNESS: unknown
180         (D) TOPOLOGY: unknown
182     (ii) MOLECULE TYPE: DNA (genomic)
186     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
188 CTAGGGAGCT CCAGGACCTT GAGGTGCACC TGGACCTTGT AATTG            45
190 (2) INFORMATION FOR SEQ ID NO: 10:
192     (i) SEQUENCE CHARACTERISTICS:
193         (A) LENGTH: 35 base pairs
194         (B) TYPE: nucleic acid
195         (C) STRANDEDNESS: unknown
196         (D) TOPOLOGY: unknown
198     (ii) MOLECULE TYPE: DNA (genomic)
202     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
204 ATTATGCATC TTGACTACCT AGGTCCAAAC CAAAC                       35
206 (2) INFORMATION FOR SEQ ID NO: 11:
208     (i) SEQUENCE CHARACTERISTICS:
209         (A) LENGTH: 66 base pairs
210         (B) TYPE: nucleic acid
211         (C) STRANDEDNESS: unknown
212         (D) TOPOLOGY: unknown
214     (ii) MOLECULE TYPE: DNA (genomic)
218     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
220 GTCATATGTT CCATCTGCAG AGCAGATCTT GGAATTCGTT AAGCAAATCT CGAGTCAGTA   60
222 ACTATA                                                      66
224 (2) INFORMATION FOR SEQ ID NO: 12:
226     (i) SEQUENCE CHARACTERISTICS:
227         (A) LENGTH: 66 base pairs
228         (B) TYPE: nucleic acid
229         (C) STRANDEDNESS: unknown
230         (D) TOPOLOGY: unknown
232     (ii) MOLECULE TYPE: DNA (genomic)
236     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
238 TATAGTTACT GACTCGAGAT TTGCTTAACG AATTCCAAGA TCTGCTCTGC AGATGGAACA   60
```

RAW SEQUENCE LISTING   DATE: 02/23/2005
PATENT APPLICATION: US/09/057,016   TIME: 15:06:06

Input Set : N:\Crf3\RULE60\09057016.raw.txt
Output Set: N:\CRF4\02232005\I057016.raw

```
240 TATGAC                                                                  66
242 (2) INFORMATION FOR SEQ ID NO: 13:
244      (i) SEQUENCE CHARACTERISTICS:
245          (A) LENGTH: 33 base pairs
246          (B) TYPE: nucleic acid
247          (C) STRANDEDNESS: unknown
248          (D) TOPOLOGY: unknown
250     (ii) MOLECULE TYPE: DNA (genomic)
254     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
256 CGACCTAGGT GATGACGTCA TAGCAATTAA CGT                                     33
258 (2) INFORMATION FOR SEQ ID NO: 14:
260      (i) SEQUENCE CHARACTERISTICS:
261          (A) LENGTH: 33 base pairs
262          (B) TYPE: nucleic acid
263          (C) STRANDEDNESS: unknown
264          (D) TOPOLOGY: unknown
266     (ii) MOLECULE TYPE: DNA (genomic)
270     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
272 TAATTGCTAT GACGTCATCA CCTAGGTCGA CGT                                     33
274 (2) INFORMATION FOR SEQ ID NO: 15:
276      (i) SEQUENCE CHARACTERISTICS:
277          (A) LENGTH: 4 amino acids
278          (B) TYPE: amino acid
279          (C) STRANDEDNESS: unknown
280          (D) TOPOLOGY: unknown
282     (ii) MOLECULE TYPE: peptide
286     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
288     Ala Gly Asp Arg
289     1
291 (2) INFORMATION FOR SEQ ID NO: 16:
293      (i) SEQUENCE CHARACTERISTICS:
294          (A) LENGTH: 510 base pairs
295          (B) TYPE: nucleic acid
296          (C) STRANDEDNESS: unknown
297          (D) TOPOLOGY: unknown
299     (ii) MOLECULE TYPE: DNA (genomic)
301     (vi) ORIGINAL SOURCE:
302          (A) ORGANISM: pBGC291 Fusion
304     (ix) FEATURE:
305          (A) NAME/KEY: CDS
306          (B) LOCATION: 1..510
309     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
311 ATG TCT TAC AGT ATC ACT ACT CCA TCT CAG TTC GTG TTC TTG TCA TCA          48
312 Met Ser Tyr Ser Ile Thr Thr Pro Ser Gln Phe Val Phe Leu Ser Ser
313 1               5                   10                  15
315 GCG TGG GCC GAC CCA ATA GAG TTA ATT AAT TTA TGT ACT AAT GCC TTA          96
316 Ala Trp Ala Asp Pro Ile Glu Leu Ile Asn Leu Cys Thr Asn Ala Leu
317             20                  25                  30
319 GGA AAT CAG TTT CAA ACA CAA CAA GCT CGA ACT GTC GTT CAA AGA CAA          144
```

```
320 Gly Asn Gln Phe Gln Thr Gln Gln Ala Arg Thr Val Val Gln Arg Gln
321     35                  40                  45
323 TTC AGT GAG GTG TGG AAA CCT TCA CCA CAA GTA ACT GTT AGG TTC CCT    192
324 Phe Ser Glu Val Trp Lys Pro Ser Pro Gln Val Thr Val Arg Phe Pro
325     50                  55                  60
327 GCA GGC GAT CGG GCT GGT GAC CGT GCA GGA GAC AGA GAC TTT AAG GTG    240
328 Ala Gly Asp Arg Ala Gly Asp Arg Ala Gly Asp Arg Asp Phe Lys Val
329 65              70              75              80
331 TAC AGG TAC AAT GCG GTA TTA GAC CCG CTA GTC ACA GCA CTG TTA GGT    288
332 Tyr Arg Tyr Asn Ala Val Leu Asp Pro Leu Val Thr Ala Leu Leu Gly
333                 85                  90                  95
335 GCA TTC GAC ACT AGA AAT AGA ATA ATA GAA GTT GAA AAT CAG GCG AAC    336
336 Ala Phe Asp Thr Arg Asn Arg Ile Ile Glu Val Glu Asn Gln Ala Asn
337             100                 105                 110
339 CCC ACG ACT GCC GAA ACG TTA GAT GCT ACT CGT AGA GTA GAC GAC GCA    384
340 Pro Thr Thr Ala Glu Thr Leu Asp Ala Thr Arg Arg Val Asp Asp Ala
341             115                 120                 125
343 ACG GTG GCC ATA AGG AGC GCG ATA AAT AAT TTA ATA GTA GAA TTG ATC    432
344 Thr Val Ala Ile Arg Ser Ala Ile Asn Asn Leu Ile Val Glu Leu Ile
345     130                 135                 140
347 AGA GGA ACC GGA TCT TAT AAT CGG AGC TCT TTC GAG AGC TCT TCT GGT    480
348 Arg Gly Thr Gly Ser Tyr Asn Arg Ser Ser Phe Glu Ser Ser Ser Gly
349 145                 150                 155                 160
351 TTG GTT TGG ACC TCT GGT CCT GCA ACT TGA                           510
352 Leu Val Trp Thr Ser Gly Pro Ala Thr
353                 165
356 (2) INFORMATION FOR SEQ ID NO: 17:
358     (i) SEQUENCE CHARACTERISTICS:
359         (A) LENGTH: 169 amino acids
360         (B) TYPE: amino acid
361         (D) TOPOLOGY: linear
363     (ii) MOLECULE TYPE: protein
365     (vi) ORIGINAL SOURCE:
366         (A) ORGANISM: pBGC291 Fusion
368     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
370 Met Ser Tyr Ser Ile Thr Thr Pro Ser Gln Phe Val Phe Leu Ser Ser
371 1               5                   10                  15
373 Ala Trp Ala Asp Pro Ile Glu Leu Ile Asn Leu Cys Thr Asn Ala Leu
374             20                  25                  30
376 Gly Asn Gln Phe Gln Thr Gln Gln Ala Arg Thr Val Val Gln Arg Gln
377     35                  40                  45
379 Phe Ser Glu Val Trp Lys Pro Ser Pro Gln Val Thr Val Arg Phe Pro
380     50                  55                  60
382 Ala Gly Asp Arg Ala Gly Asp Arg Ala Gly Asp Arg Asp Phe Lys Val
383 65              70              75              80
385 Tyr Arg Tyr Asn Ala Val Leu Asp Pro Leu Val Thr Ala Leu Leu Gly
386                 85                  90                  95
388 Ala Phe Asp Thr Arg Asn Arg Ile Ile Glu Val Glu Asn Gln Ala Asn
389             100                 105                 110
```

VERIFICATION SUMMARY   DATE: 02/23/2005
PATENT APPLICATION: US/09/057,016   TIME: 15:06:07

Input Set : **N:\Crf3\RULE60\09057016.raw.txt**
Output Set: **N:\CRF4\02232005\I057016.raw**

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:34 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:34 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vi)
L:35 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:483 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:19
L:592 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21